DOCUMENT RESUME

ED 215 965 SP 020 011

AUTHOR Tamashiro, Roy T.; And Others
TITLE Evaluating the Impact of New Priorities in Teacher Education.
PUB DATE Feb 82
NOTE 35p.; Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education (Houston, TX, February, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Attitude Change; Cooperating Teachers; Core Curriculum; Educational Trends; *Essay Tests; *Field Experience Programs; Grading; Higher Education; *Multicultural Education; *Preservice Teacher Education; *Student Attitudes; Teacher Attitudes; Urban Education

ABSTRACT

The results of three evaluation studies, which examined the impact of new program priorities in teacher education programs, are discussed. Program priorities evaluated by these studies were: (1) larger and earlier doses of field and clinical experiences in preservice teacher education; (2) preparation of teachers to work effectively in multicultural settings; and (3) coordination and/or standardization of preservice learning experiences. The first study, which explored the trend toward more field and clinical experiences in preservice programs, surveyed cooperating teachers involved with a core program for sophomore teacher candidates. Results suggested that cooperating teachers believe that more practice is better for prospective teachers. Cooperating teachers also endorsed the idea that these field experiences should be structured and well supervised. The second study examined the attitudinal changes in teacher candidates who participated in a program which included a multicultural emphasis and urban and suburban placements. These settings provided school populations of diverse cultural backgrounds. Responses to a pre- and post-experience questionnaire indicated that teacher candidates had a slight increase to more favorable attitudes toward inner city experiences. It is tentatively concluded that significant changes in attitudes develop slowly. The third study evaluated whether instructors in two teacher preparation courses graded common essay tests consistently. Instructors graded essays of students in their own and other teachers' sections. Wide discrepancies were found in the grading despite explicitly outlined grading criteria. It is suggested that risks of inconsistently graded essay tests pose a threat to their integrity as exit examinations. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED215965

Evaluating the Impact of New Priorities in Teacher Education*

Roy T. Tamashiro, Webster College

(Session Organizer)

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Roy T. Tamashiro

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

*A Thematic Session presentation at the 1982 Annual Meeting of the American Association of College for Teacher Education (AACTE), February 19, 1982, Houston, Texas.

Running head: Evaluating new priorities

SP 020 011

Title: "Evaluating the Impact of New Priorities in Teacher Education"

Presenters: Dr. Roy T. Tamashiro (Session organizer)
Webster College, St. Louis, MO

Dr. Richard C. Callahan
The Ohio State University, Columbus, OH

Dr. Alma Adams
Bennett College, Greensboro, NC

Dr. Marilynne Snook
University of Vermont, Burlington, VT

SUMMARY

New priorities and directives in teacher education often require colleges and universities to reorganize their programs without full knowledge of the positive and negative consequences or impacts on those involved, especially the teacher candidates themselves. The purposes of this session were: 1) to discuss the results of three evaluation studies which examined the impact of new program priorities in teacher education programs; and (2) to outline recommendations for future practices and procedures based on the results of these studies.

Aspects of three priorities in teacher education were considered in this session: One current trend is toward including larger doses of field and clinical experiences in preservice teacher education programs, especially field-based instruction and providing more actual instructional responsibilities for teacher candidates early in their program. Cooperating teachers critically review one program's change toward this priority in the first study reported in this session. Another current priority is to prepare teachers to work effectively in multicultural settings. The second study focuses on this priority by examining the attitudinal changes in teacher candidates who participated in a program which included a multicultural emphasis and urban and suburban field placements. Another priority involves the effort to improve the quality of teacher education programs and courses by attempting to coordinate and/or standardize the learning experiences. The third study reported in this session evaluates one possible practice resulting from this priority; namely, the administration of program or course examinations which are the same for all teacher candidates, regardless of instructors.

The first study involved a new (three-year old) undergraduate teacher education program in which teacher candidates in a multidisciplinary introduction to education course sequence devoted 30% of their course time in elementary and secondary schools. The teacher candidates assumed instructional responsibilities at several points in the two-quarter program. The 92 cooperating teachers who were involved in the program participated in this survey-type evaluation study. Questions focused on the desirability, feasibility, and satisfaction of the highly-involved field activities. The

results show strong concurrence among teachers that involving more time and greater responsibility early in the teacher education program were valuable, practicable and beneficial. National and state standards involving greater field involvement seem to have been supported as both philosophically desirable and logistically feasible in this program example.

Training educators to work effectively and positively in multicultural settings has been an aspiration in many teacher education programs. The study reported here assessed the degree to which positive multicultural attitudes could be fostered by placing teacher candidates in actual multicultural settings, such as urban schools. This study involved 96 teacher candidates in a pre-post assessment comparison design. The score differences on the Attitudes Toward Inner City Questionnaire suggested that the type of field setting per se (urban vs. suburban) was not markedly associated with more favorable attitudes. There were slightly more positive attitudes among all groups in the final (post) assessment. It seemed that preservice teachers were more preoccupied with general questions such as "can I really teach?" or "will I be able to manage 30 pupils?" rather than the more specific concern involved in the multicultural priority (such as "How will I work with learners from diverse backgrounds?"). Perhaps the skills and attitudes in preparing teachers to work in multicultural settings are very complex and hence less likely to be achieved early in undergraduate teacher education programs.

The third study examined the issue of coordinated or standardized curricula and instruction in teacher education programs as a means of improving quality. Specifically it evaluated the degree of instructor uniformity in grading essay examination questions used in teacher education programs and courses. Twelve instructors who taught the same introduction to education course were involved in this study. The same curriculum outline and exams were used by all instructors in their sections. Although they graded their own exams for the course, each instructor also graded student essays from other sections for the purposes of this study. A set of criteria was explicity outlined and there was extensive discussion about grading procedures before the exam. In spite of this careful preparation, the degree of agreement and interrater correlations were low. Inconsistent grading of the essay tests were continuing despite the extensive effort involving in standardizing the grading procedures. While the priority of coordinating and standardizing curricula and instruction many not only mean standardized grading procedures, this study underscores some of the complexity and difficulty in achieving this aim.

Although the results of the studies reported in this session may not be generalizable to other efforts to implement current teacher education priorities, they can serve as models for the evaluation of those eforts. Through systematic evaluations such as these, teacher edcators can derive valuable data regarding the practicality, the successfulness, or the desirability of the attempts to address current priorities. These data can be helpful in making more informed decisions about future directions in teacher education.

Biographical Summaries:

Dr. Roy T. Tamashiro is assistant professor of education and coordinator of the social sciences concentration area, Webster College. His research in teacher education focuses on teacher education curricula, and on the developmental processes of teachers and teacher candidates.

Dr. Richard C. Callahan is assistant professor of education, The Ohio State Univesity and has been a principal developer of a major undergraduate teacher education core program. His research interests include teacher education curriculum an urban education.

Dr. Alma S. Adams, an associate professor of art at Bennett College, is involved in teacher education curriculum development, and research in multicultural education and teacher education.

Dr. Marilynne Snook, assistant professor of home economics education, University of Vermont, is involved in curriculum development and research in home economics professional teacher education and in the supervision of home economics teacher education.

Evaluating the Impact of New Priorities

in Teacher Education

CONTENTS

Summary	2
Biographical Summary of Presenters	4
Study #1: Cooperating Teacher's Views of Early Field Experiences for Teacher Candidates	5
Study #2: Developing Multicultural Attitudes Among Teacher Candidates in Core Courses	15
Study #3: Reliability of Essay Tests Used in Teacher Education Programs	24

STUDY #1: Cooperating Teacher's Views of Early Field Experiences for Teacher Candidates

Richard C. Callahan and Roy T. Tamashiro

In teacher education, there is a belief that the more practical experience one has a preservice teacher, the more likely one is to become an effective teacher. Traditionally the period of practice occurs in a culminating student teaching or apprentice teaching experience, but many teacher educators believe the even more preservice practical experience is necessary. Evidence of this is seen in some accreditation boards and state boards requiring more field-based practice time by teacher education programs seeking accreditation.

In spite of the common sense wisdom of the value of practice experiences, it is not clear what kinds of practice experiences in the various phases of the teacher candidates' programs are most beneficial. Also it is not clear what role cooperating teachers should play when working with novice teacher candidates.

This study was an initial effort to evaluate the "more practice earlier dictum". The evaluation was an informal survey of cooperating teachers who had worked with sophomore teacher candidates in a state university's education core program.

The teacher candidates spent about 40 hours in elementary or secondary schools during the two-quarter program. About one fourth of this time was devoted to classroom observations, such as the charting of pupil-teacher

interaction, or the assessment of pupil reading level. Each teacher candidate was paired with a peer-partner during these field experiences, so that they could observe each other, provide mutual support and check perceptions. The remainder of the field time involved the sophomores teacher candidates in actual instruction. The first teaching epsiode occured three weeks into the course, the second five weeks later for two days, an the third about eight weeks later for five days.

The field experiences were followed by debriefing sessions in which the feedback from cooperating teachers, peers and the college instructor was reviewed.

From the beginning, the teacher candidates, and college instructors were pleased about the field experiences. Teacher candidates felt being in schools made the topics discussed more relevant and realistic, while instructors felt the school involvement helped motivate and keep student interest high.

In spite of the favorable reception on the university campus, it was not clear whether the elementary and secondary school cooperating teachers also felt positive about the sophomore level program. Did they feel it was appropriate, beneficial and helpful to include teacher candidates in their classrooms in instructional experiences at such an introductory level? The developers of the program examined this question by conducting a survey of cooperating teachers who participated in the program during the third year of the program. Eighty-four teachers (38 elementary and 58 secondary) responded to a questionnaire near the end of the academic year. This was approximately an 88% return rate. These teachers had participated as cooperating teachers in the program for at least three consecutive quarters and as such, provided a mature viewpoint, reflecting knowledge and experience of the changes that had

occurred within the program over this period.

The instrument used in the survey contained both objective and open-ended questions. Initial questions asked for general reactions to the core program's field component. Eighty-five percent of the cooperating teachers voiced interest in supervising another student in the core program. When asked about the best means of communicating program information, 77% of the cooperating teachers recommended an on-site orientation session. Coupled with written program information, 81% of the teachers suggested that the teacher candidates visit the school and their cooperating teacher before their actual classroom assignments, and that the teacher candidates should use the telephone for follow-up communications with the teacher.

Even though the nature of field activities were unlike traditional student teacher placements, 95% of the cooperating teachers who responded found the core program's objective clear, and 98% of them felt they clearly understood their responsibilities. All but one of the teachers thought their assigned responsibilities were appropriate.

Whereas the extent of the cooperating teachers' involvement in the "student teaching" experience has become understood through tradition at the college, the extent of cooperating teachers' involvement with teacher candidates at the beginning stages of the teacher education program was undefined. The questionnaire asked cooperating teachers to what extent they should be involved in the teacher candidates' lessons. Table 1-1 summarizes the teachers' responses. Most of the teachers (71%) felt they should suggest a topic for the teacher candidate to teach, and 71% of them indicated they wanted to be involved in more than merely approving the topics. The majority

of teachers (61%) thought their primary role should be to give feedback after the teacher candidate had taught the lesson. Less than half of the teachers thought they should be involved in the details of lesson planning (36%) or in the total planning-teaching-evaluating process (39%). Three of the 84 respondents thought cooperating teachers should not be involved with the preparation process at all. Overall the pattern of responses suggest an emerging consensus among cooperating teachers that when working with beginning teacher candidates they should be involved in suggesting topics to teach and providing feedback on lesson effectiveness. There is noticeable disagreement about any further involvment by the cooperating teachers.

Insert Table 1-1 about here

The questionnaire asked cooperating teachers about some how desirable various practices were in field placements for sophmore teacher candidates. Their responses are summarized in Table 1-2. Seventy-four teachers (88%) thought they should be able to approve or disapprove of who gets placed in their classrooms. At the time, the program assigned teacher candidates to teachers who volunteer to act as cooperating teachers.

Teachers did not endorse program's practice of pairing student teachers as shown in the third item.

Insert Table 1-2 about here

The 40 hours of field experience, for the sophomore teacher candidates,

were less than desirable for 80% of the teachers.

This course drew teacher candidates from all program areas (elementary education, science education, physical education, etc.) who eventually expected to teach at various levels from preschool through grade 12. This diversity was viewed as valuable and stimulating for discussion and exposure to a variety of perspectives, but teacher candidates sometimes felt awkward when assigned to a grade level or subject area different from their future specialty program. More than half (58%) of the teachers recommended that the teacher candidates be placed in classrooms of the age group and specialty area of their training (see Table 1-2, Item #5). This issue seemed to be open to further debate and discussion. All but one teacher recommended that there be a designated time for cooperating teachers and teacher candidates to discuss their field experience together. Teacher candidates were scheduled to be in the schools only when the cooperating teachers' classes were in session.

In all, the cooperating teachers favored each of the proposed alternative procedures for sophomore field placements. However, since the rationales and consequences for both current practices and the proposed alternatives were not elicited, it is not clear whether the cooperating teachers would still endorse these recommendations in the context of other practices in their school.

The cooperating teachers also had the opportunity to write comments and recommendations throughout the questionnaire. Sixty-five of the 84 teachers wrote at least one additional comment beyond the objective items on the questionnaire. There were 26 different comments, suggestions, recommendations and reactions. Table 1-3 lists the seven most frequent recommendations for program improvement found among the respondents. The first recommendation corroborates results in Table 1-2, with 20 teachers saying they need more time

to talk with the teacher candidates. The second most frequent recommendation, that teacher candidates have more opportunity to observe before actually doing instruction, was a new idea suggested by 11 teachers. Nine teachers repeated the suggestion in Table 1-2 that teacher candidates be assigned to their major areas only and seven teachers suggested that the forms used to evaluate teacher candidates be changed. This latter idea, along with the fifth (more time needed for the cooperating teacher to talk with the university instructor) and seventh (the university instructors need to observe the teacher candidate more) recommendations were new.

Insert Table 1-3 About here

It should be noted that Table 1-3 lists only the teachers' comments which could be understood as recommendations. There were 36 teachers who freely commented about the general value and excellence of the program.

In summary, the survey of cooperating teachers involved with a core program for sophomore teacher candidates suggests that the cooperating teachers do agree with the idea that more practice earlier is better for those who want to become teachers. They are in general agreement with teacher educators who see immense value in placing teacher candidates in school classrooms early in their programs. The cooperating teachers also endorsed the idea that field experiences should be structured and well-supervised at this level. Most of their recommendations involved increasing the supervision by various means: more discussion time, better evaluation forms, more contact with university instructor, and minimal number of teacher candidates per teacher. This survey was informative to the developers of the core program in

measuring cooperating teachers' endorsement of the fundamental concepts of the program and gathering additional recommendations helpful in shaping the future of the field component of the program.

While this study evaluated one example of an attempt to shape early field experiences for teacher candidates, the findings may be valuable to other teacher educators who are trying to redesign field activities for teacher candidates. It is likely that cooperating teachers everywhere will be supportive of more extensive field activities, including classroom instruction, by teacher candidates very early on in their preparation programs.

Table 1-1

Extent of Desirable Involvement with Sophomore Teacher Candidates

N = 84

Type of Involvement	Cooperating Teachers Should Be Involved at This Level		Cooperating Teachers Should Not Be Involved at This Level	
	N	%	N	%
Cooperating teacher suggests a topic.	60	71%	24	29%
Teacher only approves the topic.	22	26%	62	74%
Teacher gives feedback on the lesson's effectiveness.	51	61%	33	39%
Teacher assists in developing objectives, matrials, & methods.	30	36%	54	64%
Teacher involved in all phases of planning, implementation and evaluation of lessons.	33	39%	51	61%

Table 1-2

Desriability of Various Cooperating Teacher Roles

N = 84

Role	Desirable N / %	Undesirable N / %	No Response N / %
1. Cooperating teacher should know their role before participation.	83 99%	1 1%	0 0%
2. Teacher should be able to approve or veto teacher candidate placed in his/her class.	74 88%	7 8%	3 4%
3. Teacher should be assigned no more than one teacher candidate.	77 92%	5 6%	2 2%
4. Field experience time should be increased.	67 80%	10 12%	7 8%
5. Sophomore field experiences should be limited to teacher candidates' specialty areas.	49 58%	25 30%	10 12%
6. A structured period for teacher and teacher candidate to discuss the experiences is needed.	83 99%	0 0%	1 1%

Table 1-3

Recommendations for Program Improvement

Recommendations		f
1.	More time is needed for debriefing of field experiences between the cooperating teacher and the teacher candidate.	20
2.	The teacher candidate should observe the classroom they will teach in more prior to large group instruction.	11
3.	The teacher candidate should have field experiences only in their area of preparation.	9
4.	The forms used by the cooperting teacher to evaluate the teacher candidate need to be improved.	7
5.	The cooperating teacher and University supervisor need to discuss the program and the progress the teacher candidate is making more often.	4
6.	Teacher candidates need even more field experiences.	4
7.	The University supervisor needs to observe the teacher candidate in the teaching act more.	4

STUDY #2: Developing Multicultural Attitudes Among Teacher Candidates in Core Courses

Richard C. Callahan and Roy T. Tamashiro

Teacher educators now realize that providing educational equity for the large number of diverse cultural groups in American society requires more than the addition or modification of existing courses such as ethnic history of minority studies.[1] Combating discrimination and offering equal educational opportunities require that teacher training institutions make significant structural and organizational changes.[2] This article describes one attempt at restructuring a university's teacher education program to prepare teachers for working in multicultural contexts.

The teacher educators in this college assumed that the cultivation of knowledge and skills for a multicultural perspective required a redesign of the teacher education program from beginning (freshman) to end (senior or post-degree certification). The initial restructuring created an integrated core-foundations program which replaced separate courses in educational psychology, general teaching methods, and social foundations of education. While the review and revisions of other aspects of the teacher education programs (methods courses, student teaching, etc.) continue, the new core program has been instituted and has been operating for more than four years.

The core program has been a two-quarter experience for sophomores, involving one-half of their load (6 credit hours) per quarter. In this program, teacher candidates are introduced to child development, topics in

educational psychology, general teaching methods, human relation skills, cultural diversity, sociology of education and similar foundations topics in field and laboratory based experiences. There were no more than 24 students in each section of the core program. All sections used the same curriculum outline and the same set of textbooks. The field assignments were uniform, although each section is assigned to a different school. It was felt that such standardization of the curriculum provided a standard and more predictable base for the students' specialty areas (elementary education, English education, music education, physical education, etc.). These certification programs can rely on the core program providing a common set of skills, knowledge and experiences for their students.

Multicultural values and concepts are emphasized throughout the core program. The concepts and values stressed by the instructors include: heterogeneous thinking, combating untrue or harmful stereotypes, valuing deversity and uniqueness, and the emphasis on consensus and negotitation rather than conformity.[3] Teacher candidates were exposed to these concepts and values via teacher modeling, class debriefing and role playing. These values were further enhanced by the natural diversity of student beliefs in each section, represented by a cross-section of majors, socioeconomic levels, and personal experiences. This environment promoted sensitivity, heterogeneous thinking and negotiation rather than conformity. In addition, a two-week unit was devoted to diversity among learners in which the issues, concepts and values involved in multicultural education are explicity faced in both classroom discussion and laboratory work.

The field placement was viewed as the most impactful experience which engendered multicultural attitudes in the core program. Teacher candidates

were placed in at least two different types of school settings in which the pupil populations varied significantly, during their two quarters. These included urban, rural and suburban schools. The field experiences included activities such as observing and identifying discipline styles of teachers, assessing pupils' cognitive and reading levels, and examining teachers' roles. The teacher candidates planned and taught a two-day unit in the first quarter and a 5-day unit in the second quarter. Over the long range, students were placed in a variety of cultural settings throughout their teacher preparation programs. Issues of race, sex, ethnicity, learning disabilities and other dimensions of learner diversity were faced in the field sites and then discussed in class after each school visit. The instructors helped teacher candidates clarify their beliefs on issues such as stereotyping, prejudice, discrimination, self-fulfilling prophesies, and the processes for providing personal and pupil satisfaction in teaching in these various settings.

Given the core program's efforts to develop rudimentary skills and positive values in working in multicultural settings, it seemed important to evaluate whether the teacher candidates were progressing toward the development of these skills and values, so that subsequent courses and programs could build on this foundation. Were teacher candidates in the core program developing more positive multicultural attitudes? A study assessing the teacher candidates' attitudes toward diversity was conducted to answer this question. It was assumed that more "favorable" attitudes toward diverse populations were important elements in equipping teachers to work in multicultural settings. A further aim of this study was to compare attitude

change between teacher candidates placed in urban (inner city) schools and those placed in suburban schools.

The Inner City Questionnaire[4] which identifies the degree to which individuals had favorable attitudes toward inner city children, schools and environment, was chosen for use with the teacher candidates involved in the core program. Levine and Feeney report reliability and validity levels of the instrument to be acceptable.[5] The questionnaire was expanded to 51 items to reflect additional attitudes and feelings about teaching strategies needed in the inner city, characteristics of the teaching environment, personal attributes of inner city teachers and self-reflective items.

The questionnaire was administered to 78 teacher candidates in the core program. Twenty-five male and 53 female teacher candidates, between the ages of 18-24 participated in the study. None of the teacher candidates had previous field experiences in urban schools. Thirty-six teacher candidates in this sample were assigned to suburban schools, while 42 were in inner-city (predominantly Black) schools. Some of the teacher candidates (14 assigned to the inner city and 10 assigned to suburban schools) completed the Inner City Questionnaire at the beginning of the core program and again 20 weeks later. Others (28 assigned to the inner city and 26 assigned to the suburban schools) answered the questionnaire at the end of the program only.

Table 2-1 summarizes the favorable, unfavorable and neutral responses on the Inner City Questionnaire for each of the four groups. In all groups there was a plurality or majority of responses which were favorable rather than unfavorable or netural. The number and percentages of favorable attitudes in the two groups tested twice also increased slightly in the second testing. When the students assigned to inner city schools were compared to the students

assigned to suburban schools, there was not a marked difference in the percentage of favorable responses on the questionnaire. Also, the percentages of favorable responses between the groups tested at the end of the program only and the groups tested twice were not noticeably different.

Insert Table 2-1 about here

The change over time among the items was significant (p. 05) overall in groups I nd III in 10 of the 51 items. Of these, 8 were changes toward more favorable attitudes and 2 were toward less favorable attitudes. In the group assigned to the innder city, the change was significant in 4 items, 2 toward more favorable and 2 toward less favorable. In the group assigned to the suburban schools, 6 items changed significantly, 5 toward more favorable attitudes, and 1 toward less favorable attitude. By these figures, there appears to be no important differences in teacher candidates' attitudes whether they were placed in an inner city or in a suburban school for the core experience.

This evaluation of teacher candidates' attitudes in the core program revealed: (1) that there was a slight increase to more favorable attitudes toward inner city experiences in the course of the program; and (2) that there was not an appreciable difference between those students placed in an inner city school and those placed in suburban schools.

These results were illuminating in that there appeared to be very little change in attitude and very little attitudinal differences between groups regarding multicultural experiences in the course of the two-quarter core

experience (20 instructional weeks). This was at first distressing to the program developers because more noticeable changes on the attitude measure were expected to result from the intensity and the multiplicity of strategies to develop multicultural values in the program. On the other hand, there was a noticeable lack of marked changes over time and the lack of differences between groups. The changes were in the positive direction but slight. The resistance to changes in attitudes seemed to be the dominant feature among the teacher candidates in this study. Yet this resistance was consistent with the common-sense belief that attitudinal change occurs very slowly and only after a long period of time in which the new attitudes are sytematically supported.

Teacher candidates placed in the inner city schools were intuitively expected to finish the core program with more favorable attitudes on the Inner City Questionnaire than the teacher candidates who were placed in suburban schools. In actuality the type of school in which they were placed did not matter. Favorable attitudes prevailed nearly equally among teacher candidates placed in both types of schools. It appeared that other factors such as instructor modeling, class discussion, content of the courses and other characteristics in the common core experience ittself rather than the major characteristic of the school site per se were responsible for the small shifts toward more favorable attitudes. One can speculate that teacher candidates at the sophomore core experience level are mainly preoccupied with discovering whether they themselves can work successfully with students in a classroom at all, and the question of whether they can work with diversity among the learners and with multicultural concerns is much more remote to them. Thus the inner city classroom versus the suburban classroom has similar meaning for

the teacher candidate at this point, namely the opportunity to see if he or she can teach in the classroom. Perhaps placement in an inner city school or a suburban school would make a more noticeable difference later in the teacher candidates' program when they are more psychologically prepared to address issues of diversity.

In all, the evaluation study underscored the following ideas when preparing teacher candidates to work effectively toward positive multicultural attitudes and abilities:

(1) The values and skills of working with diverse learners need to be presented to teacher candidates using a variety of methods and experiences over a long period of time.

(2) Since attitude change occurs only slightly in the short term (20 weeks), teacher educators should aim for consistent, cumulative changes in attitude over the entire period of the teacher candidates' program.

(3) Placing teacher candidates in diverse settings (such as an inner city school) for the purpose of developing multicultural attitudes is only mininally valuable during teacher candidates' early training period.

(4) Although the type of school placement in the core program may not matter, experiences in actual school classrooms, including teaching experiences during the core program, do seem to be the foundation for the development of multicultural values.

NOTES

[1]James A. Banks, Teaching Strategies for Ethnic Studies, Boston: Allyn and Bacon, 1979

Geneva Gay, "Currriculum for Multicultural Education," in Frank H. Klassen and Donna M. Gollnick (editors), Pluralism and the American Teacher, Washington D.C. Ethnic Heritage for Teacher Education, 1977.

Alfred Lightfoot, Urban education in Social Perspective, Chicago: Rand McNally, 1978.

[2]Standards for Accreditation of Teacher Education Washington D.C.: National Council for Accreditation of Teacher Education, 1977.

[3]Virginia Cornea Jones. "Characteristics of Multicultural Education," The Teacher Educator, Vol. 12, 1977 (4 Spring), pp. 26-27.

[4]Murray Levine, F. Dunn, S. Brochinsky, J. Bradley, & K. Donlan "Project Scranton: Service and Training in the Clinical Experience of Student Teachers. Progress Report" New Haven, Conn.: Psycho-Educational Clinic, Yale University, 1968 (mimeographed).

[5]Murray Levine and George M. Feeney, "The Effects of Practice Teaching in Inner City Schools on Attitudes Toward Teaching in Inner City Schools," Notre Dame Journal of Education, Fall 1971, pp. 244-255.

Table 1

Group Atitudes Toward Inner City Experiences

n = 78

Teacher Candidate Groups	Number/% of Items at Time 1			Number/% of Items at Time 2		
	Favorable	Unfavorable	Neutral	Favorable	Unfavorable	Neutral
I. Assigned to Inner City School n = 14	27/53%	10/20%	14/27%	31/61%	5/10%	15/29%
II. Assigned to Inner City School n = 28	*	*	*	24/47%	11/22%	16/31%
III. Assigned to Suburban School n = 10	25/49%	7/14%	19/37%	33/64%	10/20%	8/16%
IV. Assigned to Suburban School n = 26	*	*	*	30/59%	7/14%	14/27%

*Note: Groups II and IV were tested at Time 2 only.

Study #3: Reliability of Essay Tests Used in Teacher Education Programs

Roy T. Tamashiro, Alma Adams and Marilynne Snook

The growing need to assure quality and competence among graduates of teacher education programs has stimulated renewed interest in various forms of uniform exit examinations and criteria. The essay examination has been popular with educators because it is relatively easy to prepare, and provides an excellent sampling of students' real learning, according to its advocates (1;3). However, the essay format has also been strongly criticized for its subjectivity in grading and its lack of reliability (1;3). Is this debate relevant to teacher educators contemplating the use of essay items on exit examinations? This article discusses a study which evaluated whether the instructors in two teacher preparation courses graded common essay tests in a consistent manner, and it reviews the implications of the study for teacher preparation programs.

Background

Numerous studies have illustrated the subjectivity and lack of reliability involved in the grading of essay tests (4; 5; 6; 7; 9; 11). They confirm Stalnaker's(10) early conclusion that "the typical essay test as typically handled. . . is not reliably graded and, therefore, cannot stand alone as a good measuring instrument (p.499)".

Various suggestions have been made to increase the reliability of essay tests. For example, Payne(8) suggests that the number of questions be increased and the extensivity demanded by each question minimized. The Educational Testing Service(1973) recommends that the question be stated in enough detail so that students will answer it in terms of the desired aspect

rather than covering a variety of topics. Another factor which seems to account for grader variance is the lack of guidelines of clearly defined criteria for the graders to follow. Content outlines were provided in several studies while in other studies only the number of total possible points was given to the graders.

While these suggestions appear sound, essay tests which incorporate these recommendations have not been evaluated for their reliability. In this study, several of these recommendations were used in constructing the essay tests in an attempt to increase grading reliability. First, the essay questions were limited in scope (i.e., minimized extensivity); second, the questions were pilot-tested with former students to check that the questions would be answered according to the desired aspect. Third, specific grading criteria were outlined for each essay question, and discussions were held among instructors (the graders) regarding the interpretation of the questions, the grading procedures and criteria.

The Study

Essay questions from two introductory-level education courses in a large Midwestern university were used in this study. One essay question from the final examination given in each of the two courses was selected for this study. The essay item was identical for students in all the various sections of each course. In the first course (Course A), the essay item was as follows:

> "Describe the three goals of helping suggested by Carkhuff and Gazda. For each goal define and describe the process by which the goal can be attained."

In the second course (Course B), the essay item selected for study was

the following:

> One method of discouraging disruptive classroom behavior is to give "penalty points" to individual students when they violate classroom rules and to detain them from recess, lunch or other designated time. Which of the three teacher perspectives (Learning Theorists, Humanistic, Cognitive) would say this is an effective means of classroom management? (b) Using the assumptions of the perspective you chose, give two reasons why teachers who have this perspective would say this is an effective method.

A set of guidelines for assigning grades or points and a sample answer to these questions were used by the instructors.

Seven instructors from Course A and five instructors from Course B participated in the study. They graded the essays of students who were in their own sections and also of those from the other sections. In the grading of essays from other sections, the students' names, section instructors' names, and the grades previously assigned were masked. The essay item in Course A had a maximum possible score of 20 points. The range of scores assigned to the 128 student essays was 0-20 (mean = 15.6, S.D. = 3.30). The instructors' degree of agreement and interrater correlations (r) are summarized in Table 3-1. Overall, instructor pairs gave identical scores 28% of the time and were within 2 points of each other (10% error range) 62% of the time.

INSERT TABLE 3-1 ABOUT HERE

In the second course, a letter grading system was used with a maximum grade of 4.0 (A). Fractions of points were also assignable. The range of student scores here was 0.0 to 4.0 (mean = 3.3; S.D. = .71) among 93 essays. Table 3-2 summarizes the degree of rater agreement and the interrater correlations among instructors in Course B. Overall perfect agreement among pairs of instructors was 20% and 32% agreement within 1/3 of a letter grade (as in B to B+). None of the interrater correlations was statistically significant in Course B.

INSERT TABLE 3-2 ABOUT HERE

Discussion

In order to maximize reliability among the instructors, the recommendations of previous investigators were incorporated. These recommendations included (1)restricting the scope of the questions, (2)pilot testing the questions, and (3)providing instructors with specific grading criteria and sample answers. Despite this effort, the degree of grader agreement was only slightly better in Course A than grader agreements in previous reliability studies of essay tests in which the above suggestions had not been used. In Course B, the rater agreement was either comparable to or poorer than that of previous studies. The instructors' percent agreement and correlations in Course A can be interpreted as marginally acceptable (except for Instructor #4A). In Course B, the percent agreement and correlations were not within the acceptable range. It is questionable whether teacher educators would want to invest as much effort specifying essay questions, pilot-testing them and developing elaborate guidelines and grading criteria if these efforts

do not improve the grading to an acceptable level.

Although the study results do not pinpoint the sources of discrepancy among the raters, it is possible to conclude that the specific suggestions which were followed, by themselves, do not consistently improve the degree of agreement among graders. The magnitude of difference in grader agreement between Course A and Course B suggests that there were other influences on whether graders were in agreement. In reviewing the differences between the essay questions in the two courses, the question for Course A appears more specific, finite and restricted in scope than the Course B question. The Course B question appears to require more analysis and application by the student. The grading criteria for Course B were more difficult for the instructors to apply. These differences in the formulation of the questions and the complexity of applying the grading criteria are possible sources of the discrepancy. Often these problems are only identified after the pilot testing has been completed, and corrections cannot be made until the test is given in another class. However, until these discrepancies are identified and corrected, it is generally not wise for teacher educators to assume that essay items on common exit examinations of courses or programs will be graded consistently, even when careful precautions such as those described above are taken.

Although we may not yet have confidence that essay tests can be graded reliably, teacher educators may continue to experiment with various ways of formulating the questions, the grading criteria or the sample answers. We may also experiment with different ways of administering or grading essays which might improve grader consistency. For example, one instructor grades all student responses for question #1 from the several sections; a second

instructor grades all the question #2 responses, and so forth.

In sum, while the merits of essays should warrant including essay items in common exit examinations for teacher education courses or programs, the risks of inconsistently grading them remains a threat to the integrity of evaluating teacher candidates. Continued caution and further field testing of essay grading practices are needed before we can use essays with confidence as an acceptable components of common exit criteria.

References

1. Coffman, W.E. "Essay Examinations" In R.L. Thorndike (Ed.) Educational Measurement (2nd edition), Washington, D.C.: American Council on Education, 1971.
2. Ebel, R.L. Essentials of Educational Measurement (3rd edition), Englewood Cliffs, N.J.: Prentice-Hall, 1979.
3. Educational Testing Service, Multiple Choice Questions: A Close Look, Princeton, N.J.: Educational Testing Service, 1973.
4. Eells, W.C. "Reliability of Repeated Grading of Essay Type Questions," Journal of Educational Psychology, 1930, 21, 48-52.
5. Harding, L.H. "An Illustration of the Effects of Subjectivity Upon Evaluation of Essay Type Responses." In L.H. Harding, Testing, Failing, Cramming, Cheating, Passing, Columbus, Ohio: Association for the Study of Educology, 1977.
6. Marshall, J.C. and Powers, J.M. "Writing Neatness, Composition Errors, and Essay Grades" In D.A. Payne and R.F. McMorris(Ed.) Educational and Psychological Measurements(2nd edition), Morristown, N.J.: General Learning Press, 1975.
7. Myers, A.E., McConville, C. and Coffman, W.E., "Simplex Structures in the Grading of Essay Tests." Educational and Psychological Measurement, 1966, 26, 41-54.

8. Payne, D.A., The Specification and Measurement of Learning Outcomes,

Waltham, Mass.: Blaisdell Publishing, 1968.

9. Ruch, G.M., The Objective or New-Type Examination, Chicago: Scott, Foresman, 1929, pp. 78-81.

10. Stalnaker, J.M. "The Essay Type Examination", Educational Measurements, Washington, D.C.: American Council on Education, 1951.

11. Starch, D. and Elliot, E.C., "Reliability of Grading High School Work in English," School Review, 1912, 20, 442-457.

TABLE 1

Grader Agreements in Course A

Instructor	% Perfect Agreement	% Agreement Within 2 Points	r
#1A	15%	45%	.91*
#2A	28%	76%	.71*
#3A	15%	45%	.57*
#4A	20%	55%	-.75*
#5A	35%	80%	.89*
#6A	50%	68%	.93*
#7A	16%	55%	.73*
OVERALL	28%	62%	.84*

*p .01

**p .001

TABLE 2

Grader Agreements in Course B

Instructor	% Perfect Agreement	% Agreement Within 1/3 Grade	r
#1B	60%	60%	.22
#2B	12%	22%	-.34
#3B	20%	31%	-.26
#4B	0%	30%	.33
#5B	15%	18%	-.05
OVERALL	20%	32%	-.07